

WHAT IS CLAIMED IS:

- 1 537
2 A
3 1. A computer-implemented method of displaying a document using a
4 browser, the method comprising:
5 accessing the document;
6 searching the document to identify text patterns in the document which are
7 relevant to a plurality of concepts;
8 marking locations of the text patterns in the document; and
9 displaying the document using the browser such that the text patterns in
10 the document which are relevant to the plurality of concepts are annotated.
- 1 2. The method of claim 1 wherein the browser is an Internet Explorer
2 browser and the searching and marking use information about the document stored in a
3 Document Object Model configured by the Internet Explorer browser.
- 1 3. The method of claim 1 wherein searching the document comprises:
2 determining text patterns associated with the plurality of concepts; and
3 searching the document to identify text patterns in the document which
4 match the text patterns associated with the plurality of concepts.
- 1 4. The method of claim 1 wherein:
2 the browser is an Internet Explorer browser; and
3 the searching and marking is performed using methods provided by an
4 IHTMLTxtRange interface.
- 1 5. The method of claim 1 wherein:
2 the browser is an Internet Explorer browser; and
3 the searching and marking is performed using methods provided by an
4 IMarkupServices interface.
- 1 6. The method of claim 1 wherein marking the locations of the text
2 patterns in the document comprises inserting annotation tags in front of and after each
3 text pattern identified in the document to be relevant to a concept from the plurality of
4 concepts, wherein the annotation tags for each text pattern identify the concept to which
5 the text pattern is relevant.

1 7. The method of claim 6 wherein displaying the document
2 comprises:
3 for each text pattern identified as being relevant to a concept from the
4 plurality of concepts:
5 determining the concept to which the text pattern is relevant from
6 the annotation tags surrounding the text pattern;
7 determining style information to be used for annotating the text
8 pattern, wherein the style information is associated with the concept to which the text
9 pattern is relevant; and
10 annotating the text pattern based on the style information.

1 8. The method of claim 7 wherein annotating the text pattern
2 comprises highlighting the text pattern in a color indicated by the style information.

1 9. The method of claim 1 further comprising:
2 calculating a score for each concept in the plurality of concepts, the score
3 indicating relevance of the document to the concept; and
4 displaying a relevance indicator for each concept in the plurality of
5 concepts based on the score for the concept.

1 10. The method of claim 9 wherein calculating the score for each
2 concept in the plurality of concepts comprises:
3 for each concept in the plurality of concepts:
4 determining frequency of text patterns in the document identified
5 as being relevant to the concept; and
6 calculating the score based on the frequency of the text patterns.

1 11. A computer-implemented method of displaying a multi-page
2 document using a browser, the method comprising:
3 accessing a multi-page document;
4 displaying a section of the multi-page document in a first viewing area of a
5 display;
6 determining information about the contents of the multi-page document;
7 displaying a thumbnail image in a second viewing area of the display
8 based on the contents information determined from the multi-page document, the

9 thumbnail image displaying the contents of the multi-page document in a continuous
10 form; and
11 emphasizing an area of the thumbnail image corresponding to the section
12 of the multi-page document displayed in the first viewing area.

1 12. The method of claim 11 wherein:
2 determining information about the contents of the multi-page document
3 comprises:
4 determining dimension information for the contents; and
5 determining coordinate information for the contents; and
6 displaying the thumbnail image comprises:
7 displaying the contents in the thumbnail image based on the
8 dimension and coordinate information for the contents.

1 13. The method of claim 12 wherein displaying the contents in the
2 thumbnail image based on the dimension and coordinate information for the contents
3 comprises:
4 for each content:
5 determining position of the content in the thumbnail image by
6 dividing the coordinate and dimension information for the content by a reduction ratio.

1 14. The method of claim 11 wherein determining information about the
2 contents of the multi-page document comprises:
3 determining information about text entities contained in the multi-page
4 document;
5 determining dimension and coordinate information for the text entities;
6 determining if the text entities are relevant to a plurality of concepts; and
7 if a text entity is relevant to a concept in the plurality of concepts:
8 associating the text entity with style information for the concept,
9 the style information indicating a manner of annotating text entities in the multi-page
10 document which are relevant to the concept.

1 15. The method of claim 14 wherein displaying the thumbnail image
2 comprises:

3 if a text entity is relevant to a concept in the plurality of concepts:
4 displaying the text entity in the thumbnail image using the style
5 information for the concept.

1 16. The method of claim 15 further comprising:
2 modifying the style information for a concept;
3 in response to the modification:

4 identifying text entities in the multi-page document which are
5 relevant to the concept; and
6 dynamically changing the display of the identified text entities in
7 the thumbnail image based on the modified style information.

1 17. The method of claim 11 wherein determining information about the
2 contents of the multi-page document comprises:

3 determining information about forms contained in the multi-page
4 document; and
5 determining dimension and coordinate information for the forms.

1 18. The method of claim 11 wherein determining information about the
2 contents of the multi-page document comprises:

3 determining information about image elements contained in the multi-page
4 document; and
5 determining dimension and coordinate information for the image elements.

1 19. A computer-implemented method of displaying a multi-page
2 document using a browser, the method comprising:

3 accessing the multi-page document;
4 searching the multi-page document to identify text patterns in the
5 document which are relevant to a plurality of concepts;
6 marking locations of the text patterns in the document;
7 displaying a section of the multi-page document in a first viewing area of a
8 display such that the text patterns in the multi-page document which are relevant to a
9 plurality of concepts are annotated;

10 determining information about the contents of the multi-page document;
11 displaying a thumbnail image in a second viewing area of the display
12 based on the contents information determined from the multi-page document, the
13 thumbnail image displaying the contents of the multi-page document in a continuous
14 form; and
15 emphasizing an area of the thumbnail image corresponding to the section
16 of the multi-page document displayed in the first viewing area.

20. A system for displaying a document using a browser, the system comprising:

- a processor; and
- a memory coupled to the processor and configured to store a plurality of modules for execution by the processor, the plurality of modules including:
 - a module for accessing the document;
 - a module for searching the document to identify text patterns in the document which are relevant to a plurality of concepts;
 - a module for marking locations of the text patterns in the document; and
 - a module for displaying the document using the browser such that the text patterns in the document which are relevant to the plurality of concepts are annotated.

1 21. The system of claim 20 wherein the browser is an Internet Explorer
2 browser and the module for searching and the module for marking use information about
3 the document stored in a Document Object Model configured by the Internet Explorer
4 browser.

1 22. The system of claim 20 wherein the module for searching the
2 document further comprises:
3 a module for determining text patterns associated with the plurality of
4 concepts; and
5 a module for searching the document to identify text patterns in the
6 document which match the text patterns associated with the plurality of concepts.

1 23. The system of claim 20 wherein:
2 the browser is an Internet Explorer browser; and
3 the modules for searching and marking use methods provided by an
4 IHTMLTxtRange interface.

1 24. The system of claim 20 wherein:
2 the browser is an Internet Explorer browser; and
3 the modules for searching and marking use methods provided by an
4 IMarkupServices interface.

1 25. The system of claim 20 wherein the module for marking the
2 locations of the text patterns in the document comprises a module for inserting annotation
3 tags in front of and after each text pattern identified in the document to be relevant to a
4 concept from the plurality of concepts, wherein the annotation tags for each text pattern
5 identify the concept to which the text pattern is relevant.

1 26. The system of claim 25 wherein the module for displaying the
2 document comprises:
3 for each text pattern identified as being relevant to a concept from the
4 plurality of concepts:
5 a module for determining the concept to which the text pattern is
6 relevant from the annotation tags surrounding the text pattern;
7 a module for determining style information to be used for
8 annotating the text pattern, wherein the style information is associated with the concept to
9 which the text pattern is relevant; and
10 a module for annotating the text pattern based on the style
11 information.

1 27. The system of claim 26 wherein the module for annotating the text
2 pattern comprises a module for highlighting the text pattern in a color indicated by the
3 style information.

1 31. The system of claim 30 wherein:
2 the module for determining information about the contents of the multi-
3 page document comprises:
4 a module for determining dimension information for the contents;
5 and
6 a module for determining coordinate information for the contents;
7 and
8 the module for displaying the thumbnail image comprises:
9 a module for displaying the contents in the thumbnail image based
10 on the dimension and coordinate information for the contents.

1 32. The system of claim 31 wherein the module for displaying the
2 contents in the thumbnail image based on the dimension and coordinate information for
3 the contents comprises:
4 for each content:
5 a module for determining position of the content in the thumbnail
6 image by dividing the coordinate and dimension information for the content by a
7 reduction ratio.

1 33. The system of claim 30 wherein the module for determining
2 information about the contents of the multi-page document comprises:
3 a module for determining information about text entities contained in the
4 multi-page document;
5 a module for determining dimension and coordinate information for the
6 text entities;
7 a module for determining if the text entities are relevant to a plurality of
8 concepts; and
9 if a text entity is relevant to a concept in the plurality of concepts:
10 a module for associating the text entity with style information for
11 the concept, the style information indicating a manner of annotating text entities in the
12 multi-page document which are relevant to the concept.

1 34. The system of claim 33 wherein the module for displaying the
2 thumbnail image comprises:

3 if a text entity is relevant to a concept in the plurality of concepts:
4 a module for displaying the text entity in the thumbnail image
5 using the style information for the concept.

1 35. The system of claim 34 wherein the plurality of modules stored in
2 the memory further comprise:

3 a module for modifying the style information for a concept;
4 in response to the modification:
5 a module for identifying text entities in the multi-page document
6 which are relevant to the concept; and
7 a module for dynamically changing the display of the identified
8 text entities in the thumbnail image based on the modified style information.

1 36. The system of claim 30 wherein the module for determining
2 information about the contents of the multi-page document comprises:

3 a module for determining information about forms contained in the multi-
4 page document; and
5 a module for determining dimension and coordinate information for the
6 forms.

1 37. The system of claim 30 wherein the module for determining
2 information about the contents of the multi-page document comprises:

3 a module for determining information about image elements contained in
4 the multi-page document; and
5 a module for determining dimension and coordinate information for the
6 image elements.

1 38. A system for displaying a multi-page document using a browser,
2 the system comprising:

3 a processor; and
4 a memory coupled to the processor and configured to store a plurality of
5 modules for execution by the processor, the plurality of modules module including:
6 a module for accessing the multi-page document;

00636030-080900

7 a module for searching the multi-page document to identify text
8 patterns in the document which are relevant to a plurality of concepts;

9 a module for marking locations of the text patterns in the
10 document;

11 a module for displaying a section of the multi-page document in a
12 first viewing area of a display such that the text patterns in the multi-page document
13 which are relevant to a plurality of concepts are annotated;

14 a module for determining information about the contents of the
15 multi-page document;

16 a module for displaying a thumbnail image in a second viewing
17 area of the display based on the contents information determined from the multi-page
18 document, the thumbnail image displaying the contents of the multi-page document in a
19 continuous form; and

20 a module for emphasizing an area of the thumbnail image
21 corresponding to the section of the multi-page document displayed in the first viewing
22 area.

1 39. A computer program product stored on a computer readable
2 storage medium for displaying a document using a browser, the computer program
3 product comprising:

4 code for receiving accessing the document;

5 code for searching the document to identify text patterns in the document
6 which are relevant to a plurality of concepts;

7 code for marking locations of the text patterns in the document; and

8 code for displaying the document using the browser such that the text
9 patterns in the document which are relevant to the plurality of concepts are annotated.

1 40. A computer program product stored on a computer readable
2 storage medium for displaying a multi-page document using a browser, the computer
3 program product comprising:

4 code for accessing a multi-page document;

5 code for displaying a section of the multi-page document in a first viewing
6 area of a display;

7 code for determining information about the contents of the multi-page
8 document;
9 code for displaying a thumbnail image in a second viewing area of the
10 display based on the contents information determined from the multi-page document, the
11 thumbnail image displaying the contents of the multi-page document in a continuous
12 form; and
13 code for emphasizing an area of the thumbnail image corresponding to the
14 section of the multi-page document displayed in the first viewing area.

1 41. A computer program product stored on a computer readable
2 storage medium for displaying a multi-page document using a browser, the computer
3 program product comprising:
4 code for accessing the multi-page document;
5 code for searching the multi-page document to identify text patterns in the
6 document which are relevant to a plurality of concepts;
7 code for marking locations of the text patterns in the document;
8 code for displaying a section of the multi-page document in a first viewing
9 area of a display such that the text patterns in the multi-page document which are relevant
10 to a plurality of concepts are annotated;
11 code for determining information about the contents of the multi-page
12 document;
13 code for displaying a thumbnail image in a second viewing area of the
14 display based on the contents information determined from the multi-page document, the
15 thumbnail image displaying the contents of the multi-page document in a continuous
16 form; and
17 code for emphasizing an area of the thumbnail image corresponding to the
18 section of the multi-page document displayed in the first viewing area.